CONSENSUS ALGORITHMS OVER FADING CHANNELS

Kevin Chan*, Ananthram Swami*, Qing Zhao†, and Anna Scaglione†

*Army Research Laboratory, Adelphi, MD
†University of California at Davis, CA
(kevin.s.chan, ananthram.swami)@us.army.mil, qzhao@ucdavis.edu, ascaglione@ucdavis.edu

ABSTRACT

Consensus algorithms permit the computation of global statistics via local communications and without centralized control. We extend previous results by taking into account fading and unidirectional links in ring and random 2-D topologies. We study conditions for convergence and present simulation results to verify the analytical results in this paper. We compare the performance of consensus algorithms with a tree-based (centralized) approach. Additionally, we implement a slotted ALOHA protocol and compare its performance to that under the initial assumption of perfect scheduling.

I. INTRODUCTION

In the classical distributed estimation problem, sensors make noisy observations of a scalar. The (weighted) average of these observations is a good estimate under many scenarios; this average may be computed by nodes sending their estimates to a fusion center (the centralized version), or by exchange of information between nodes. The basic consensus problem addresses the convergence of such schemes. Centralized approaches require the computation of a tree, rooted at the fusion node, aggregation of data up the tree, and dissemination of the consensus value down the tree. See [1] for a discussion of issues and approaches related to scalable aggregation and scalable inference. Re-computation of the tree would be required under mobility, duty-cycling and fading. The role of the fusion node may need to be rotated in order to avoid single points of failure, and depletion of resources at the nodes neighboring the fusion node. Decentralized approaches may be preferable under these circumstances.

The distributed consensus problem is also of interest in several military contexts. Wireless networks are becoming more prevalent in tactical scenarios. Further, decentralized algorithms are desired so that resource requirements are distributed throughout the network. Olfati-Saber *et al* [3] briefly describe several military-relevant applications that distributed consensus would improve. Flocking theory and formation control techniques are applied to unmanned vehicle platforms for obstacle avoidance and waypoint finding tasks. With consensus, each agent or node would make heading or velocity updates without centralized control. Consensus in distributed sensing environments would decentralize the requirements of communication and computational costs. And consensus has been used as a mechanism for decentralized network-wide time synchronization.

Detailed discussion of the state-of-the-art may be found in Section III. In the basic setup, nodes $i = 1, \cdots, n$, have initial estimates or measurements $x_i(0)$, and the problem is to obtain a common weighted average $\sum_i a_i x_i(0)$ via decentralized exchange of messages. Implicit is the notion of local exchanges in a one-hop neighborhood (that, for example, may be defined by an SNR threshold criterion on the received signal). In the literature the analysis is often carried out by assuming that the underlying connectivity matrix, or the neighborhood set is static, and, for wireless applications, a random geometric graph is assumed [2]. Similar to [6], we consider a broadcast update, but unlike it,

we explicitly consider a fading channel and take into account collisions. Coupled with an SNR threshold criterion, this leads to a random neighborhood set. Under this model, the connectivity graph is not *symmetric* in any given snapshot (node i may reliably receive node j's estimate but not vice versa), and the consensus algorithm is not sum-preserving. We consider the convergence of the consensus algorithm under various assumptions on the fading. We also investigate an Aloha type random access scheme and compare performance with tree-based approaches.

Our main contributions are in studying the effect of fading and collisions on the performance of wireless consensus gossiping and in comparing its cost (measured in terms of number of transmitted packets or convergence time) with that of a tree-based approach.

II. MODEL AND ASSUMPTIONS

We consider a network of n nodes with locations denoted by v_i, $i = 1, ..., n$. We consider fading, path loss and additive noise. The fading process, $g(t)$, is modeled as complex circularly symmetric noise with zero mean and variance $\sigma_g^2 = 1$. Let $s_i(t)$ denote the unit amplitude signal transmitted by node i, and P_o the fixed transmitted power. Then the signal received at node k from node i in time slot t can be written as

$$y_k(t) = g_{ki}(t)\sqrt{\frac{P_o}{d_{ik}^\alpha}}s_i(t) + \nu_k(t)$$

where $g_{ki}(t)$ denotes the fading coefficient associated with the transmission from node i to node k, $\nu_k(t)$ is AWGN with common variance σ_ν^2, $d_{ik} = d_{ki}$ is the distance between nodes i and k, and α denotes the path loss exponent.

We assume that a transmission can be successfully decoded if the SNR exceeds a threshold SNR_T. Let d_o denote the range below which this threshold SNR is exceeded in a non-fading channel, i.e.,

$$\text{SNR}_T = \frac{P_o}{\sigma_\nu^2 d_o^\alpha} \ .$$

In the presence of fading, the effective transmission range r_g, i.e., the range at which the SNR threshold is exceeded, is random, due to the randomness of the fading $g(\cdot)$. Now, node i can decode node j's transmission successfully if

$$\text{SNR}_{ij} := P_o|g_{ij}|^2/\sigma_\nu^2 d_{ij}^\alpha \geq \ \text{SNR}_T,$$

i.e., if $|g_{ij}|^2 d_{ij}^{-\alpha} \geq d_o^{-\alpha}$. When $g(t)$ is complex Gaussian (Rayleigh fading) $|g|^2$ is exponentially distributed with parameter 1, the probability that node j can successfully talk to node i is

$$P_{ij} = \text{Pr}(\ \text{SNR}_{ij} \geq \ \text{SNR}_T) = \exp\left\{-\left(\frac{d_{ij}}{d_o}\right)^\alpha\right\} \ .$$

Note that this does not take collisions into account.

Example 1: Ring Topology. If n nodes are uniformly distributed on a ring with radius r, then

$$d_{ij} = 2r \sin\left(\frac{\pi|i-j|}{n}\right) \ .$$

Report Documentation Page

1. REPORT DATE **OCT 2010**	2. REPORT TYPE	3. DATES COVERED **00-00-2010 to 00-00-2010**

4. TITLE AND SUBTITLE **Consensus Algorithms Over Fading Channels**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Army Research Laboratory,Adelphi,MD,20783**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
in Proceedings of IEEE Military Communication Conference (MILCOM), 31 Oct ? 3 Nov, 2010, San Jose, CA

14. ABSTRACT

Consensus algorithms permit the computation of global statistics via local communications and without centralized control. We extend previous results by taking into account fading and unidirectional links in ring and random 2-D topologies. We study conditions for convergence and present simulation results to verify the analytical results in this paper. We compare the performance of consensus algorithms with a tree-based (centralized) approach. Additionally, we implement a slotted ALOHA protocol and compare its performance to that under the initial assumption of perfect scheduling.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**6**	

Further, with $m_o = \frac{n}{\pi} \sin^{-1}(\frac{d_o}{2r})$, where d_o is the nominal transmission range,

$$\Pr(\ \mathrm{SNR}_{ij} \geq \ \mathrm{SNR}_T) = \exp\left\{ -\left(\frac{\sin(\pi(i-j)/n)}{\sin(\pi m_o/n)} \right)^{\alpha} \right\} .$$

For a given fade value, $|g|$, the effective range is given implicitly by

$$m = \arg \max_k d_{0k}, \text{s.t. } |g|^2 \left[\frac{d(m_o)}{d_{0k}} \right]^{\alpha} \geq 1 .$$

II-A. The basic consensus algorithm

At the end of one *round* of transmissions, all nodes will update their estimates, as

$$x_i(t+1) = [1 - m_i(t)h]x_i(t) + h \sum_{j \in \mathcal{N}_i(t)} x_j(t) \qquad (1)$$

where t is discrete time, h is a coupling parameter, $\mathcal{N}_i(t)$ is the set of nodes that successfully transmit to node i in round t, and $m_i(t) = |\mathcal{N}_i(t)|$ is the *in-degree* of node i. The parameter h is a measure of stubbornness or stiffness: Smaller values of h reflect less confidence in neighbors' estimates and lead to slow network convergence. On the other hand, large values can lead to oscillations and lack of convergence. We can allow the stubbornness factor to be dependent on both the transmitting and receiving nodes, h_{ij}:

$$x_i(t+1) = [1 - \mu \sum_{j \in \mathcal{N}_i(t)} h_{ij}]x_i(t) + \mu \sum_{j \in \mathcal{N}_i(t)} h_{ij}x_j(t)$$

which reduces to (1) when $\mu h_{ij} = h, \forall i, j$. Note that the protocol is synchronous, but the randomness in the network update is induced by packets that are dropped due to fading or collisions.

In matrix form

$$\mathbf{x}(t+1) = \mathbf{W}(t)\mathbf{x}(t) \qquad (2)$$

where matrix $\mathbf{W}(t)$ is defined by

$$W_{ij}(t) = \begin{cases} 1 - \mu \sum_{j \in \mathcal{N}_i(t)} h_{ij} & i = j \\ \mu h_{ij} & j \in \mathcal{N}_i(t) \\ 0 & \text{o.w.} \end{cases} \qquad (3)$$

Let $\mathbf{A}(t)$ denote the adjacency matrix, $\mathbf{D}(t)$ the diagonal matrix of in-degrees, and $\mathbf{L}(t) := \mathbf{D}(t) - \mathbf{A}(t)$, the graph Laplacian. If $h_{ij} = h, \forall i, j$, then

$$\mathbf{W}(t) = \mathbf{I} - h\mathbf{L}(t) .$$

If links are bidirectional, instantaneous disagreement can be represented in terms of the Laplacian, as shown in [3], [4],

$$\mathbf{x}'(t)\mathbf{L}(t)\mathbf{x}(t) = \sum_{i=1}^{n} \sum_{j \in \mathcal{N}_i(t)} |x_i(t) - x_j(t)|^2 .$$

Thus, one can think of CA as a steepest descent algorithm. We do not assume that links are bidirectional; hence neither $\mathbf{A}(t)$ nor $\mathbf{L}(t)$ is assumed to be symmetric.

Note that this formulation does not preclude simultaneous transmissions or receptions; collisions resulting in decoding failures will simply not be used in the update equation: the corresponding $W_{ij} = 0$.

II-B. Assumptions on the Fading

We will consider the performance of our consensus algorithm in (1), which we will call CA, under one of two different assumptions on the fading:
(**A1**) All transmissions in a given round suffer the same fade; fading is independent from round to round.
(**A2**) All fades are iid across transmitters and across rounds.

Assumption A1 is reasonable if the nodes are clustered and the time for a full round of transmission is smaller than the channel coherence time. Assumption A2 is reasonable when the nodes are dispersed or the coherence time is small compared with the duration of one round of transmissions. Note that we have not assumed any specific distribution (such as Rayleigh) for the fading under A1-A2.

We do not need explicit assumptions on scheduling; however, in this section, we assume that there are no collisions if nodes transmit simultaneously. This may be achieved if nodes have multi-packet reception capability , or proper scheduling is performed. We illustrate the impact of collisions (in an Aloha-based scheme) in Example 6 in Section VII.

Our CA in (1) is not sum preserving:

$$\sum_{i=1}^{n} x_i(t+1) = \sum_{j=1}^{n} x_j(t) \sum_{i=1}^{n} W_{ij}(t)$$
$$= \sum_{j} x_j(t)[1 - hm_j(t) + hn_j(t)]$$

where $n_j(t)$ is the number of nodes that can successfully hear j in round t, i.e., the *out-degree* of node j, and recall that $m_j(t)$ is the in-degree. In general, $m_i(t) \neq n_i(t)$, although clearly $\sum_i m_i(t) = \sum_i n_i(t), \forall t$. Under assumption A1, the fading is fully correlated (i.e., all transmissions in a given round experience the same fade); hence $m_i(t) = m(t) = n_i(t)$; and CA is sum-preserving. Under A2, the $m_i(t)$ and $n_i(t)$ are independent random variables. From the update equation in (1), we can conclude immediately that $\mathbf{W}(t)\mathbf{1} = \mathbf{1}$; i.e., $\mathbf{1}$, the $n \times 1$ vector of ones, is always a right eigenvector of $\mathbf{W}$. However, $\mathbf{1}'\mathbf{W}(t) \neq \mathbf{1}'$ unless the CA is sum-preserving. The connectivity graph is not regular (i.e., all nodes do not have the same degree), and the in-degree does not equal the out-degree, except under A1. Further note that communication links are not assumed to be symmetric under A2.

III. RELATED WORK

There has been a resurgence of interest in characterizing consensus and gossip algorithms following the seminal work of Xiao and Boyd [5], in which a randomly selected pair of nodes exchange and update their estimates at any given time slot. Conditions for convergence in mean and mean-square are derived in [5]. In Aysal *et al* [6], at a given instant, a single node broadcasts its estimates and all nodes in its one-hop neighborhood update their estimates; conditions for convergence in mean and mean square are derived, and update parameters are optimized. Instantaneous consensus suffices when the focus is not on parameter estimation *per se* but on arriving at consensus in some protocol: e.g., clock synchronization, allocation of TDMA slots or hop frequencies. Barbarossa *et al* [7] consider the consensus problem on a regular ring topology and derive convergence conditions. Vanka *et al* [8] also consider the problem in a ring topology and in a 2-D topology. In the preceding papers, a static AWGN channel is assumed. The average consensus problem under iid link and node failures has been studied by Barooah *et al* [9], [10]. Wang and Elia [11] establish conditions for convergence when the link drops are iid and delays are link-dependent.

If the communication links are static, the update algorithm is typically written as $\mathbf{x}(t+1) = \mathbf{W}\mathbf{x}(t)$ (see eq. 2), where $\mathbf{x}(t)$ is the vector of sensor estimates at time t. It is reasonable to assume that in the absence of fading, the underlying graph $\mathcal{G}_o$ and consensus algorithm $\mathbf{W}$ achieve consensus; if $\mathbf{W}$ is doubly stochastic

(which does not demand symmetry), then average consensus can be achieved. Now fading causes link drops in $\mathcal{G}_o$, yielding $\mathcal{G}(t)$ and hence $\mathbf{W}(t)$. Our algorithm in (2) is the biased compensation method of Fagnani-Zampieri [12]. However, their results are not applicable since they explicitly invoke the assumption that fades are iid in time (at each t, $G(t)$ is drawn independently) and across links; the latter assumption is equivalent to assuming that fades between any pair of nodes is independent of all others. Pereira *et al* [13] consider a special case of [12] (the underlying graph is fully connected; the iid fade process yields an Erdos-Renyi graph); their weight matrix is $\mathbf{W}(t) = I - h\mathbf{L}(t)$, where $\mathbf{L}(t)$ is the graph Laplacian. Different from [12], they assume that the initial measurement vector, $\mathbf{x}(0)$, is a random vector whose elements are iid; they derive an expression for average MSE as a function of iteration index k, and optimize the value of the coupling coefficient h. Because of the iid link activation, $E\{\mathbf{W}(t)\}$ is symmetric and doubly stochastic. Link drop models are also considered in [14]; here the link drops are independent; the drop probability can be link dependent, but links are assumed to be symmetric.

Noisy receptions over random topologies are studied in [17]. Here, the coupling coefficients (the weights for "others-information") are time-varying, L_2-summable but not L_1-summable. This persistence condition is required to ensure that the estimate is consistent. The link failure model is general: the Laplacian is iid in time and the mean Laplacian has non-zero second eigenvalue. Link failures can be correlated in space. A signal fading model rather than a link drop model is investigated in [18]. Here, the received signal is modeled as $y(t) = gx(t) + \nu(t)$, i.e., a fading additive noise channel. The fades are assumed known to the receiver and $x(t)$ is binary (0,1). Consensus to the initial *majority* decision is sought.

We consider an SNR threshold model: a link is 'on' if the received SNR exceeds a prescribed threshold. The design for a nominal AWGN channel is assumed to yield a connected graph $\mathcal{G}_o$ on which consensus can be achieved (although note that what is important is that the effective graph be connected: fading could make a disconnected graph connected sufficiently often). Fading induces link drops which are correlated; the link drop probability is link dependent. We assume that the link drops are iid in time. We do not assume the weight matrix $\mathbf{W}(t)$ or its mean to be symmetric.

Consensus over fading channels and the issue of deterministic vs. random schedules has been considered in [19] and [20]. We consider an Aloha based approach and compare performance to that of tree-based approaches.

Our model is different from [12], [13], [16], as we do not assume that link drops are iid (across links) or have the same probability of drop for every link. (If the nominal design range is d_o and path loss factor is α, then a transmission to a node d away is successful, i.e., the link is on, if $|g|^2 > (d/d_o)^\alpha$; thus the link drop probability depends upon the link. Further the link drops associated with a given transmitter are correlated; if the link to a node at distance d_1 is off, links to all nodes farther than d_1 are off; and if the link to a node d_2 away is on, links to all nodes closer than d_2 are on. Even if the fading coefficient $g(\cdot)$ associated with each link is iid, the probability of a link drop would be link dependent as in [14]). Different from [13] we assume that the initial observation vector is deterministic. Different from [14], [15], we do not assume symmetric links in $G(t)$. Different from [17], our weighting of others information does not decay with time and does not satisfy the square-summability condition. Different from [18] the initial node variables are not restricted to be binary. Our model generalizes those in [6], [7] to include fading.

IV. CONVERGENCE TO INITIAL MEAN VALUE

Let $\overline{\mathbf{W}} := E\{\mathbf{W}(t)\}$ where the expectation is wrt the fading process, $g(t)$, equivalently, the effective range r_g. Now, from (3), we obtain

$$\overline{W}_{ij} := E\{W_{ij}(t)\} = hP_{ij}, \quad i \neq j \tag{4}$$

where P_{ij} is the probability that node j can successfully transmit to node i. Let $\mathcal{I}_{ij}$ denote the indicator for successful transmission from j to i. Then, for $i = j$,

$$
\begin{aligned}
W_{ii}(t) &= 1 - hm_i(t) = 1 - h\sum_j \mathcal{I}_{ij}(t) \\
\overline{W}_{ii}(t) &= 1 - h\sum_j E\{\mathcal{I}_{ij}\} \\
&= 1 - h\sum_{j \neq i} P_{ij} := 1 - h\overline{m}_i \tag{5}
\end{aligned}
$$

where $\overline{m}_i$ is the expected number of updates received by node i (its mean in-degree).

Note that $\mathbf{W}$ and $\overline{\mathbf{W}}$ are right stochastic (row sums are unity). If the channel between every pair of nodes is stochastically reciprocal, i.e., $P_{ij} = P_{ji}$, $\forall i, j$, then $\overline{\mathbf{W}}$ is symmetric (see (4)) and hence doubly stochastic, i.e., $\sum_k \overline{W}_{kj} = \sum_k \overline{W}_{ik} = 1$, $\forall i, j$. If h is chosen such that $1 - h\overline{m}_i > 0$, $\forall i$, and if the fading process is such that $P_{ij} > 0$, $\forall j \neq i$, then all the entries of $\overline{\mathbf{W}}$ are positive. The former can be assured by choosing $h < 1/n$. The Perron-Frobenius theorem then applies, and we can conclude immediately that the maximum eigenvalue of $\overline{\mathbf{W}}$ is unique and is given by 1, and that the absolute value of the other eigenvalues are all less than unity. We can verify that $\mathbf{1}$, the $n \times 1$ vector of 1's is the eigenvector corresponding to $\lambda = 1$. Let $\mathbf{J} := \frac{1}{n}\mathbf{1}\mathbf{1}'$. It follows that the eigenvalues of $\overline{\mathbf{W}} - \mathbf{J}$ satisfy[1]

$$|\lambda(\overline{\mathbf{W}} - \mathbf{J})| < 1, \forall k \text{ . or } \rho(\overline{\mathbf{W}} - \mathbf{J}) < 1$$

where $\rho(\cdot)$ denotes the spectral radius.

If $\overline{\mathbf{W}}$ is symmetric, i.e., $P_{ij} = P_{ji}$, then the conditions $\overline{\mathbf{W}}\mathbf{1} = \mathbf{1}, \mathbf{1}'\overline{\mathbf{W}} = \mathbf{1}', \rho(\overline{\mathbf{W}} - \mathbf{J}) < 1$ hold and suffice to ensure that [5],[6, Prop 1]

$$E\left\{\lim_{t \to \infty} \mathbf{x}(t)\right\} = \mathbf{J}\mathbf{x}(0)$$

Remarks:

1) We have only made mild assumptions about the fading. To ensure that the entries of $\overline{\mathbf{W}}$ are positive, we assumed that $P_{ij} > 0$, $\forall i \neq j$. If this condition does not hold, then the entries of $\overline{\mathbf{W}}$ are only non-negative. The spectral radius of $\overline{\mathbf{W}}$ is still unity, $\lambda_1 = 1$[2], but $|\lambda| = 1$ may be a repeated root, which means that we cannot claim $\rho(\overline{\mathbf{W}} - \mathbf{J}) < 1$. Rayleigh fading satisfies the assumption $P_{ij} > 0$, as do other fading processes. This assumption implies that there is a non-zero probability that any two nodes can communicate. The assumption $P_{ij} > 0$, $\forall i \neq j$ is sufficient, but not necessary, as illustrated by Example 3 in Section VII.
2) We also assumed that $P_{ij} = P_{ji}$ i.e., the fading is stochastically symmetric. We illustrate via Example 4 in Section VII that this assumption is not necessary.
3) We did not need to invoke any of the assumptions A1-A2 on the joint distribution of the fading, i.e., fading may be iid partially or fully correlated.
4) Under the stochastic reciprocity assumption, $P_{ij} = P_{ji}$ which is satisfied by our model under (A2), $\overline{\mathbf{W}}$ is symmetric, even though $\mathbf{W}(t)$ need not be symmetric $\forall t$. Let $\mathbf{Q}_W(t) := \prod_{k=1}^{t} \mathbf{W}(t)$. Suppose that $\mathbf{Q}(t)$ has rank 1, whence it can be written as $\mathbf{Q}(t) = \mathbf{1}\mathbf{v}'$. Since $\mathbf{W}(t)$ is stochastic $\forall t$, it follows that $\mathbf{Q}(t + 1) = \mathbf{W}(t + 1)\mathbf{Q}(t) = \mathbf{1}\mathbf{v}'$. Thus if $\mathbf{Q}(t)$ is rank one and symmetric at some t, it remains so regardless of $\mathbf{W}(t + k)$, $k > 0$. Our conjecture is that if

---

[1]Details omitted due to lack of space
[2]We will assume that the eigenvalues are ordered so that λ_k is the eigenvalue with the k-th largest absolute value

$\mathbf{W}(1)$ is symmetric, then $\mathbf{Q}(t)$ will be symmetric. We test this in Example 5 in Section VII.

Thus, we have established:

Result 1: The consensus algorithm in (1) converges to the average initial value, if $h < 1/n$ and the fading satisfies $P_{ij} = P_{ji} > 0$, $\forall i \neq j$.

V. CONVERGENCE IN SECOND MOMENT

We saw that CA converges in the mean; in this section, we will establish convergence in mean-square. First we consider convergence to the instantaneous average $\bar{\mathbf{x}}(t) = \mathbf{J}\mathbf{x}(t)$. Define the instantaneous error as $\tilde{\mathbf{x}}(t) = \mathbf{x}(t) - \bar{\mathbf{x}}(t) = [\mathbf{I} - \mathbf{J}]\mathbf{x}(t)$.

If $\lambda_{\max}\left(E\left\{\mathbf{W}(t)'(\mathbf{I} - \mathbf{J})\mathbf{W}(t)\right\}\right) < 1$, then Lemma 3 of [6] assures us that CA converges in mean-square:

$$E\{||\tilde{\mathbf{x}}(t)||_2^2\} \to 0 .$$

Our CA is sum-preserving under A1; and Lemma 3 reduces to $\lambda_2(\overline{\mathbf{W}}) < 1$, where λ_2 denotes the second-largest eigenvalue. We verified this condition in the previous section. Thus, we consider assumption A2 first. We can verify that the conditions of [6, Lemma 3] are satisfied by the matrix $\mathbf{W}$ in (3), and establish the following.

Result 2: Under A2, the consensus algorithm in (1) converges to the instantaneous average in mean square, if $h < 1/n$ and $\overline{\mathbf{Y}} := E\{\mathbf{W}'(t)\mathbf{J}\mathbf{W}(t)\}$ has full rank.

We believe that $\overline{\mathbf{Y}}$ would have full rank if the degree distributions (the distribution of $m_i(t)$), is identical for all the nodes.

VI. CENTRALIZED VS. DISTRIBUTED

Even for a static topology, we only have asymptotic convergence results. To measure the convergence time in [21] the authors introduced the ϵ-*averaging time* which, in settings similar to ours, has been shown to grow as $O(n^2)$. In contrast, one can create a tree in a decentralized fashion, pass values to the root of the tree, and then distribute the consensus value, with a finite number of exchanges. The price paid is clearly that of finding the appropriate routing tables. One should be able to characterize the expected latency for this (at least via bounds). For the case of link failures, one can compute the average wait time (number of retransmissions required) and thus bound the expected latency.

We consider a simple distributed tree algorithm similar to ones that have been used for network time synchronization in wireless sensor networks, based on the fact that nodes must first learn the local topology before a distributed tree can be found. Fading is ignored during the formation of the tree. The root node broadcasts a level 0 beacon; one hop neighbors then take turns to transmit level 1 beacons. Nodes that have already heard a beacon ignore later ones. It is clear that all nodes will take turns to transmit beacons; thus n beacons will be transmitted. Further, an ACK mechanism is required so that each parent knows how many children it has; there will be $(n - 1)$ ACKs. Thus the total number of packets for tree formation is $2n - 1$. We have ignored scheduling issues. Each node will transmit a packet to its parent that contains the average value of the incoming data and its in-degree. ACKs are assumed, and retransmits will occur until successful transmission. If p denotes the probability of link success, then the average number of retransmits is $1/p$. A total of $(n - 1)$ successful transmissions is required for aggregation at the fusion center. With $p_{\min}$ denoting the probability of success over the weakest link, an upper bound on the expected number of packets is $n + n/p_{\min}$, making the usual assumption that ACKs are perfect. The root will then compute the average, which will then be disseminated down the tree. Again ACKs are required and a node will retransmit packets until all its children can successfully decode it. The average number of packets can be upper-bounded by $n + n/p_{\min}$. Thus the average number of packets for tree formation, aggregation, and dissemination can be upper-bounded by $2n(2 + 1/p_{\min})$.

VII. SIMULATION RESULTS

Example 2: We implemented the consensus algorithm over a 2D random network, with $n = 100$ nodes, coupling parameter $h = \frac{1}{2n}$, path loss exponent $\alpha = 2$, nominal transmission range $r_0 = \{0.1, 0.3, 0.5\}$. Figure 1 shows MSE vs. #iterations for a Monte Carlo simulation over $K = 50$ topology realizations. As expected convergence is faster when the range r_0 is larger, i.e., when the network is more connected.

Example 3: In Section IV, we showed that $P_{ij} > 0$, $\forall i, j$, is a sufficient condition for convergence in the mean. However, this is not a necessary condition (see Remark 1 in Section IV). To demonstrate this, we simulated CA across a ring network. (There is no compelling reason to run a consensus algorithm on a ring network, since faster convergence to average consensus can be achieved by simple token passing to the nearest neighbor in, say, the clockwise direction. We use a ring network only to obtain insights.) Figure 2 shows the empirical log(MSE) from a Monte Carlo simulation ($K = 50$). The network consisted of nodes uniformly placed on a ring, with parameters: network size $n = 100$, update parameter $h = \frac{1}{2n}$, nominal range $m_0 = 20$, path loss exponent $\alpha = 2$ (black curve), $\alpha = 6$ (blue curve). The red curve corresponds to the case where the range is chosen randomly over $\text{Unif}[0, 20]$. This corresponds to the case where some of the P_{ij}'s are zero (see Remark 1), indicating that the condition $P_{ij} > 0$ $\forall i, j$ is not necessary. As expected, convergence is slower when the path loss exponent is larger (connectivity is weaker).

Example 4: This example verifies Remark 2, that the condition $P_{ij} = P_{ji}$, $\forall i, j$ is only sufficient, not necessary. Figure 3 shows log(MSE) vs. iteration for a 2D network with $n = 100$, $a = 3$ from a Monte Carlo simulation ($K = 50$). Connectivity within the network is determined by P, where $P_{ij} \sim Unif[0.01, 0.70]$. 20% of the off-diagonal terms (i.e., links) were turned off (set to 0). The figure shows that the consensus algorithm will converge to the initial mean value. The case where $P_{ij} \sim Unif[0.01, 0.70]$ (i.e., $P_{ij} > 0$) is also shown in Figure 3. Convergence is faster in the latter case, as expected.

Example 5: This example tests the conjecture in Remark 4, that if $\mathbf{W}(1)$ is symmetric, then convergence could be faster. We implemented the consensus algorithm over a ring topology, with $n = 101$, $\alpha = 3$, and nominal transmission range $m_0 = 10$. Figure 4 shows log(MSE) vs. number of iterations for a Monte Carlo simulation ($K = 50$). In this example, $\mathbf{W}(1)$ is forced to be symmetric ($W_{ij}(1) = W_{ji}(1)$); in subsequent rounds, unidirectional links are allowed, i.e., $\mathbf{W}(t)$ need not be symmetric. Figure 4 shows that convergence is faster than in the baseline case where symmetry is not enforced on $\mathbf{W}(1)$.

Example 6: In this example, we adopt a standard slotted Aloha protocol for medium access. As before, a SNR threshold is assumed; simultaneous transmissions may or may not lead to successful transmissions in this case. We consider the performance of the consensus algorithm with respect to p, the probability that a node chooses to broadcast in each time slot. Figure 5 shows the number of time slots necessary to achieve consensus within an error threshold τ vs. the Aloha parameter p in a 2D random network with $n = \{50, 100, 200\}$, $\alpha = 3$, $r_0 = 0.1$, $SINR_T = 2$, and $\tau = 0.01$. Also, each simulation is run for $T = 10000$ time slots or until the MSE falls below the threshold. The dashed line shows the number of packets that are successfully decoded (by at least one node); the total number of transmitted packets can be estimated as pnT; results shown are averages across $K = 100$ realizations of the topology and initial values.

For larger networks, the choice of p becomes critical and using a slightly larger value of p is better than choosing a smaller value: the benefit due to fading outweighs the loss due to collisions.

Example 7: We compare the energy consumption of the consensus algorithm with that of the tree-based approach. Figure 6 shows the total number of transmitted packets required to achieve consensus within τ error vs. the nominal communication range r_0. The # packets required for the tree-based algorithm are shown, as well as those for CA with perfect scheduling and for the Aloha-based approach. Figure 7 shows a histogram comparing the total number of transmitted packets for the case where $r_0 = 0.3$. Parameters were $p = 0.02$, $\mathrm{SINR}_T = 2$, $\tau = 0.01$, $n = 100$, $K = 100$, $T = 5000$. The total required number of transmitted packets for the consensus approach is determined by scaling the number of iterations required by the size of the network n, which results in the discrete plot entries. The apparent discrete distribution of the Aloha-based approach is due to the interval chosen for the histogram. It is seen that the tree-based approach consumes less energy than CA when the nominal range, r_0, is large enough to ensure sufficient connectivity to establish a tree. When r_0 falls below this threshold, the tree-based approach fails, but CA can still achieve consensus. As the nominal communications range increases, connectivity improves and CA requires fewer packets, and so does the Aloha-based approach. Notice that random scheduling requires about twice the number of packets as does perfect scheduling.
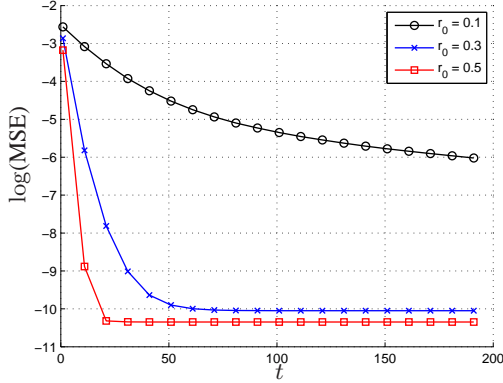

Fig. 1. Example 2: MSE vs. time for a random 2-D network; r_o is the nominal transmission range.

VIII. CONCLUSIONS

We studied the convergence of consensus algorithms when links are not symmetric. We showed that some gain can be obtained by selectively choosing bidirectional links. We also show that the link probability need not be non-zero in order for the consensus algorithm to converge. We compared the performance of an Aloha-based scheme with that under perfect scheduling, and we investigated the relative performance of consensus-based algorithms with tree-based algorithms in a fading environment. We showed that the tree-based algorithms can incur less communications cost than consensus algorithms except provided that the nominal communications range is large enough to permit the establishment of a tree. A possible extension is to use geographic gossip type protocols, which include some routing; this will probably exhibit a behavior which mediates between the robustness of consensus and the speed of the tree approach.
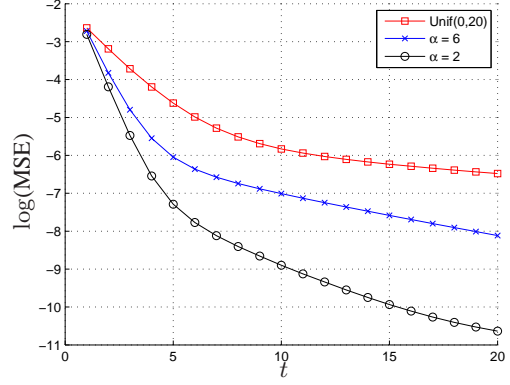

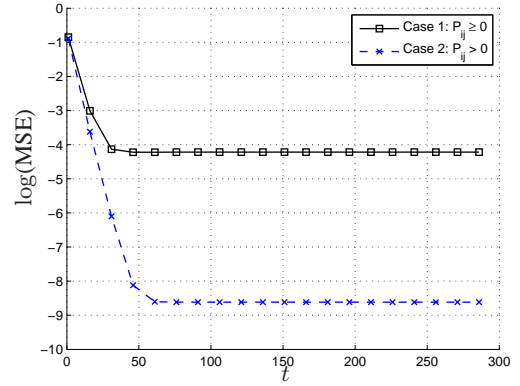**Fig. 2**. Example 3: MSE vs. time for a ring network

Fig. 3. Example 4: MSE vs. time for a 2D random network with asymmetric link probabilities P_{ij}; the solid line corresponds to the case where some of the P_{ij} are zero.
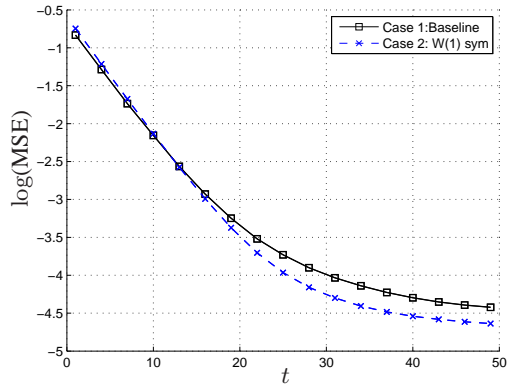

Fig. 4. Example 5: MSE vs. time for a ring network. In Case 2, only bidirectional links are used in round 1, so that $\mathbf{W}(1)$ is forced to be symmetric.
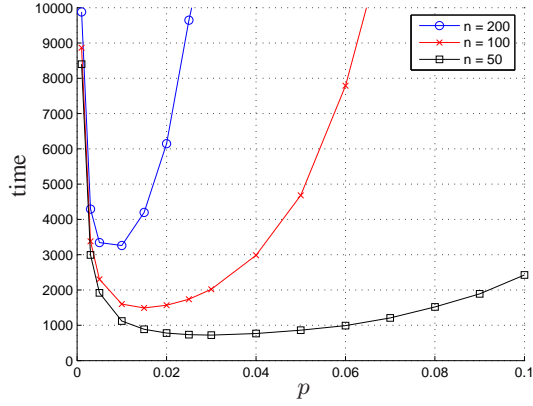
Fig. 5. Example 6: Total time vs. Aloha transmission probability p for a 2D random network; n is the number of nodes.
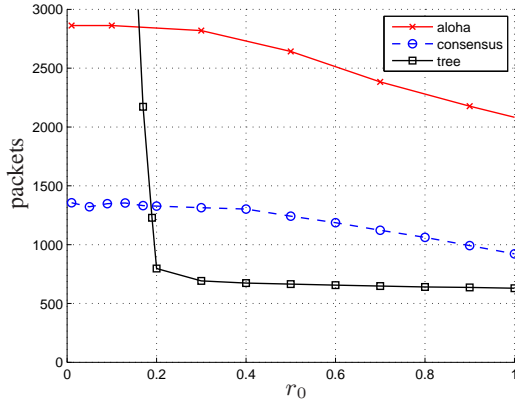

Fig. 6. Example 7: Total # packets vs. nominal range for different algorithms, for a 2D random network.
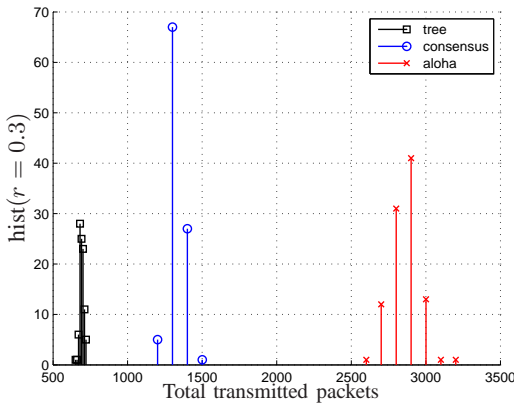

Fig. 7. Example 7: Histogram of total # packets for tree-based approach, consensus, and aloha-based.

IX. REFERENCES

[1] A. Anandkumar, A. Ephremides, A. Swami, L. Tong, "Routing for statistical inference in sensor networks", in *Handbook on Array Processing & Sensor Networks*, S. Haykin and K.J.R. Liu, Eds., IEEE-Wiley, 2010; to appear.

[2] M. Penrose, *Random Geometric Graphs*, Oxford University Press, 2003.

[3] R. Olfati-Saber, J.A. Fax, R.M. Murray, "Consensus and cooperation in networked multi-agent systems", *Proc. of the IEEE*, 95(1), 215-233, Jan 2007.

[4] S. Sardellitti, S. Barbarossa, A. Swami, "Average consensus with minimal energy consumption: optimal topology and power consumption", *Proc. EUSIPCO-10*, Aalborg, Denmark, 23-27 Aug 2010. Submitted.

[5] L. Xiao and S. Boyd, "Fast linear iteration for distributed averaging", *Sys. Contr. Lett.*, vol. 53, 65-78, 2004.

[6] T.C. Aysal, E. Yildiz, A.D. Sarwate, A. Scaglione, "Broadcast gossip algorithms for consensus", *IEEE Trans. Sig. Proc.*, 57(7), July 2009.

[7] S. Barbarossa, G. Scutari, and A. Swami, "Achieving consensus in self-organizing wireless sensor networks: the impact of network topology on energy consumption", *Proc. IEEE ICASSP*, Honolulu, April 2007

[8] S. Vanka, V. Gupta and M. Haenggi, "Power-delay analysis of consensus algorithms on wireless networks with interference", *Int'l Journal of Systems, Control and Communications*, 2009.

[9] P. Barooah, J. Hespanha, A. Swami, "On the effect of asymmetric communication on distributed time synchronization", in *Proc. IEEE Conference on Decision & Control*, New Orleans, 12-14 Dec 2007

[10] P. Barooah and A. Swami, "Recursive time synchronization in sensor networks", *Proc. IEEE MILCOM*, Nov 2008.

[11] Jing Wang, Nicola Elia, "Mean square stability of consensus over fading networks with nonhomogeneous communication delays", *Proc. IEEE CDC*, 4614-19, Dec 2008.

[12] F. Fagnani and S. Zampieri, "Average consensus with packet drop communications", *SIAM Journl of Control and Optimization*, 2009.

[13] S.S. Pereira and A. Pages-Zamora, "Fast mean-square convergence of consensus algorithms in WSNs with random topologies", *Proc. IEEE ICASSP*, 2213-2216, April 2009.

[14] S. Kar and J. Moura, "Sensor networks with random links: topology design for distributed consensus", *IEEE Trans. Sig. Proc.*, 56(7), 3315-26, July 2008.

[15] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 48, pp. 9881001, 2003.

[16] A. Y. Hatano and M. Mesbahi, "Agreement over random networks," *IEEE Transactions on Automatic Control*, vol. 50, no. 11, pp. 18671872, 2005.

[17] S. Kar and J. Moura, "Distributed consensus algorithms in sensor networks with imperfect communication: link failures and channel noise", *IEEE Trans. Sig. Proc.*, 57(1), 355-369, Jan 2009.

[18] M. Malmirchegini, Y. Ruan, Y. Mostofi, "Binary consensus over fading channels: a best affine estimation approach", *Proc. IEEE Globecom* 2008.

[19] S. Kirti and A. Scaglione, "Consensus in wireless networks with interference", *IEEE Commun. Letts.*, 2010, to appear.

[20] T.C. Aysal, A.D. Sarwate, A.G. Dimakis, "Reaching consensus in wireless networks with probabilisitc broadcast", *Proc. Allerton'09*, Allerton, Sept 2009.

[21] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms", *IEEE Trans. Info. Theory*, 52(6), 2508-2530, Jun. 2006.